

Fingerprinting with Equiangular Tight Frames

Dustin G. Mixon, Christopher J. Quinn, *Student Member, IEEE*, Negar Kiyavash, *Member, IEEE*,
and Matthew Fickus, *Member, IEEE*

Abstract—Digital fingerprinting is a framework for marking media files, such as images, music, or movies, with user-specific signatures to deter illegal distribution. Multiple users can collude to produce a forgery that can potentially overcome a fingerprinting system. This paper proposes an equiangular tight frame fingerprint design which is robust to such collusion attacks. We motivate this design by considering digital fingerprinting in terms of compressed sensing. The attack is modeled as linear averaging of multiple marked copies before adding a Gaussian noise vector. The content owner can then determine guilt by exploiting correlation between each user’s fingerprint and the forged copy. The worst-case error probability of this detection scheme is analyzed and bounded. Simulation results demonstrate the average-case performance is similar to the performance of orthogonal and simplex fingerprint designs, while accommodating several times as many users.

Index Terms—digital fingerprinting, collusion attacks, frames.

I. INTRODUCTION

DIGITAL media protection has become an important issue in recent years, as illegal distribution of licensed material has become increasingly prevalent. A number of methods have been proposed to restrict illegal distribution of media and ensure only licensed users are able to access it. One method involves cryptographic techniques, which encrypt the media before distribution. By doing this, only the users with appropriate licensed hardware or software have access; satellite TV and DVDs are two such examples. Unfortunately, cryptographic approaches are limited in that once the content is decrypted (legally or illegally), it can potentially be copied and distributed freely.

An alternate approach involves marking each copy of the media with a unique signature. The signature could be a

change in the bit sequence of the digital file or some noise-like distortion of the media. The unique signatures are called *fingerprints*, by analogy to the uniqueness of human fingerprints. With this approach, a licensed user could illegally distribute the file, only to be implicated by his fingerprint. The potential for prosecution acts as a deterrent to unauthorized distribution.

Fingerprinting can be an effective technique for inhibiting individual licensed users from distributing their copies of the media. However, fingerprinting systems are vulnerable when multiple users form a *collusion* by combining their copies to create a forged copy. This attack can reduce and distort the colluders’ individual fingerprints, making identification of any particular user difficult. Some examples of potential attacks involve comparing the bit sequences of different copies, averaging copies in the signal space, as well as introducing distortion (such as noise, rotation, or cropping).

There are two principal approaches to designing fingerprints with robustness to collusions. The first approach uses the *marking* assumption [1]: that the forged copy only differs from the colluders’ copies where the colluders’ copies are different (typically in the bit sequence). In many cases, this is a reasonable assumption because modifying other bits would likely render the file unusable (such as with software).

Boneh and Shaw proposed the first known fingerprint design that uses the marking assumption to identify a member of the collusion with high probability [1]. Boneh and Shaw’s method incorporates the results of Chor et al., who investigated how to detect users who illegally share keys for encrypted material [2]. Schaathun later showed that the Boneh-Shaw scheme is more efficient than initially thought and proposed further improvements [3]. Tardos also proposed a method with significantly shorter codelengths than those of the Boneh-Shaw procedure [4]. Several recent works investigate the relationship between the fingerprinting problem and multiple access channels, and they calculate the capacity of a “fingerprint channel” [5]–[8]. Barg and Kabatiansky also developed “parent-identifying” codes under a relaxation of the marking assumption [9], and there have been a number of other works developing special classes of binary fingerprinting codes, including [10]–[15].

The second major approach uses the *distortion* assumption. In this regime, fingerprints are noise-like distortions to the media in signal space. In order to preserve the overall quality of the media, limits are placed on the magnitude of this distortion. The content owner limits the power of the fingerprint he adds, and the collusion limits the power of the noise they add in their attack. When applying the distortion assumption, the literature typically assumes that the collusion linearly averages their individual copies to forge the host signal. Also, while results in this domain tend to accommodate fewer users than

D. G. Mixon is with the Program in Applied and Computational Mathematics, Princeton University, Princeton, New Jersey 08544 USA (e-mail: dmixon@princeton.edu).

C. J. Quinn is with the Department of Electrical and Computer Engineering, University of Illinois, Urbana, Illinois 61801 USA (e-mail: quinn7@illinois.edu).

N. Kiyavash is with the Department of Industrial and Enterprise Systems Engineering, University of Illinois, Urbana, Illinois 61801 USA (e-mail: kiyavash@illinois.edu).

M. Fickus is with the Department of Mathematics and Statistics, Air Force Institute of Technology, Wright-Patterson AFB, Ohio 45433 USA (e-mail: matthew.fickus@afit.edu).

This work was presented in part at ICASSP 2011 and SPIE 2011. The authors thank the Air Force Summer Faculty Fellowship Program for making this collaboration possible. This work was supported by NSF DMS 1042701, AFOSR F1ATA00083G004, AFOSR F1ATA00183G003, AFOSR FA9550-10-1-0345 and NRL N00173-09-1-G033. Mixon was supported by the A. B. Krongard Fellowship. Quinn was supported by the Department of Energy Computational Science Graduate Fellowship, which is provided under grant number DE-FG02-97ER25308. The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

those with the marking assumption, the distortion assumption enables a more natural embedding of the fingerprints, i.e., in the signal space.

Cox et al. introduced one of the first robust fingerprint designs under the distortion assumption [16]; the robustness was later analytically proven in [17]. Ergun et al. then showed that for any fingerprinting system, there is a tradeoff between the probabilities of successful detection and false positives imposed by a linear-average-plus-noise attack from sufficiently large collusions [18]. Specific fingerprint designs were later studied, including orthogonal fingerprints [19] and simplex fingerprints [20]. There are also some proposed methods motivated by CDMA techniques [21], [22]. Jourdas and Moulin demonstrated that a high rate can be achieved by embedding randomly concatenated, independent short codes [23]. Kiyavash and Moulin derived a lower bound on the worst-case error probabilities for designs with equal-energy fingerprints [24]. An error probability analysis of general fingerprint designs with unequal priors on user collusion is given in [25]. Some works have also investigated fingerprint performance under attack strategies other than linear averaging [26], [27] and under alternative noise models [28].

When working with distortion-type fingerprints, some embedding process is typically needed. This process is intended to make it difficult for the colluders to identify or distort fingerprints without significantly damaging the corresponding media in the process. For instance, if the identity basis is used as an orthogonal fingerprint design, then three or more colluders can easily identify and remove their fingerprints through comparisons. Indeed, different signal dimensions have different perceptual significance to the human user, and one should vary signal power strengths accordingly. There is a large body of work in this area known as *watermarking*, and this literature is relevant since fingerprinting assigns a distinct watermark to each user. As an example, one method of fingerprint embedding is known as spread spectrum watermarking. Inspired by spread spectrum communications [29], this technique rotates the fingerprint basis to be distributed across the perceptually significant dimensions of the signal [16]. This makes fingerprint removal difficult while at the same time maintaining acceptable fidelity. For an overview of watermarking media, see [30], [31].

The present paper proposes a fingerprint design under the distortion assumption. Specifically, we propose equiangular tight frames for fingerprint design and analyze their performance for the worst-case collusion. Moreover, through simulations, we show that these fingerprints perform comparably to orthogonal and simplex fingerprints on average, while accommodating several times as many users. Li and Trappe [22] also used fingerprints satisfying the Welch bound, but did not use tight frames, designed the decoder to return the whole collusion, and did not perform worst case analysis. Use of compressed sensing for fingerprint design was first suggested in [32] and [33], follow by [34]. However, [34] and [33] focused on detection schemes with Gaussian fingerprints. Additionally, [35] examined combinatorial similarities of fingerprint codes and compressed sensing measurement matrices.

We present the description of the fingerprinting problem

in Section II. In Section III, we discuss this problem from a compressed sensing viewpoint and introduce the equiangular tight frame fingerprint design. Using this design, we consider a detector which determines guilt for each user through binary hypothesis testing, using the correlation between the forged copy and the user's fingerprint as a test statistic. In Section IV, we derive bounds on the worst-case error probability for this detection scheme assuming a linear-average-plus-noise attack. Finally in Section V, we provide simulations that demonstrate the average-case performance in comparison to orthogonal and simplex fingerprint designs.

II. PROBLEM SETUP

In this section, we describe the fingerprinting problem and discuss the performance criteria we will use in this paper. We start with the model we use for the fingerprinting and attack processes.

A. Mathematical model

A content owner has a host signal that he wishes to share, but he wants to mark it with fingerprints before distributing. We view this host signal as a vector $s \in \mathbb{R}^N$, and the marked versions of this vector will be given to $M > N$ users. Specifically, the m th user is given

$$x_m = s + f_m, \quad (1)$$

where $f_m \in \mathbb{R}^N$ denotes the m th fingerprint. We assume the fingerprints have equal energy:

$$\gamma^2 := \|f_m\|^2 = ND_f, \quad (2)$$

that is, D_f denotes the average energy per dimension of each fingerprint.

We wish to design the fingerprints $\{f_m\}_{m=1}^M$ to be robust to a linear averaging attack. In particular, let $\mathcal{K} \subseteq \{1, \dots, M\}$ denote a group of users who together forge a copy of the host signal. Then their linear averaging attack is of the form

$$y = \sum_{k \in \mathcal{K}} \alpha_k (s + f_k) + \epsilon, \quad \sum_{k \in \mathcal{K}} \alpha_k = 1, \quad (3)$$

where ϵ is a noise vector introduced by the colluders. We assume ϵ is Gaussian noise with mean zero and variance $N\sigma^2$, that is, σ^2 is the noise power per dimension. The relative strength of the attack noise is measured as the *watermark-to-noise ratio* (WNR):

$$\text{WNR} := 10 \log_{10} \left(\frac{ND_f}{N\sigma^2} \right). \quad (4)$$

This is analogous to signal-to-noise ratio. See Figure 1 for a schematic of this attack model.

B. Detection

Certainly, the ultimate goal of the content owner is to detect every member in a forgery coalition. This can prove difficult in practice, though, particularly when some individuals contribute little to the forgery, with $\alpha_k \ll 1$. However, in the real world, if at least one colluder is caught, then other members could be identified through the legal process. As such, we consider

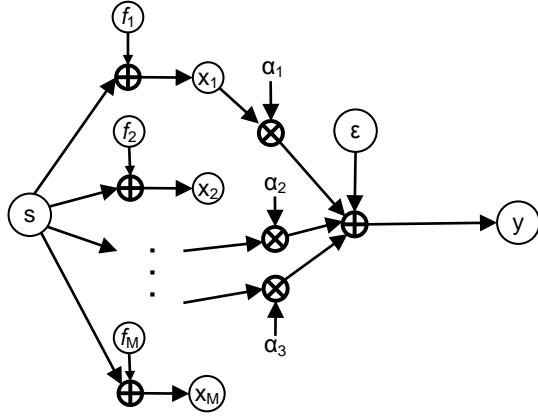


Fig. 1. The fingerprint assignment (1) and forgery attack (3) processes.

focused detection, where a test statistic is computed for each user, and we perform a binary hypothesis test for each user to decide whether that particular user is guilty.

With the cooperation of the content owner, the host signal can be subtracted from a forgery to isolate the fingerprint combination:

$$z := \sum_{k \in \mathcal{K}} \alpha_k f_k + \epsilon.$$

We refer to $\sum_{k \in \mathcal{K}} \alpha_k f_k$ as the *noiseless* fingerprint combination. The test statistic for each user m is the normalized correlation function:

$$T_m(z) := \frac{1}{\gamma^2} \langle z, f_m \rangle, \quad (5)$$

where γ^2 is the fingerprint energy (2). For each user m , let $H_1(m)$ denote the guilty hypothesis ($m \in \mathcal{K}$) and $H_0(m)$ denote the innocent hypothesis ($m \notin \mathcal{K}$). Letting τ denote a correlation threshold, we use the following detector:

$$\delta_m(\tau) := \begin{cases} H_1(m), & T_m(z) \geq \tau, \\ H_0(m), & T_m(z) < \tau. \end{cases} \quad (6)$$

To determine the effectiveness of our fingerprint design and focused detector, we will investigate the corresponding error probabilities.

C. Error analysis

Due in part to the noise that the coalition introduced to the forgery, there could be errors associated with our detection method. One type of error we can expect is the false-positive error, whose probability is denoted P_I , in which an innocent user m ($m \notin \mathcal{K}$) is found guilty ($T_m(z) \geq \tau$). This could have significant ramifications in legal proceedings, so this error probability should be kept extremely low. The other error type is the false-negative error, whose probability is denoted P_{II} , in which a guilty user ($m \in \mathcal{K}$) is found innocent ($T_m(z) < \tau$). The probabilities of these two errors depend on the fingerprints $F = \{f_m\}_{m=1}^M$, the coalition \mathcal{K} , the weights $\alpha = \{\alpha_k\}_{k \in \mathcal{K}}$, the user m , and the threshold τ :

$$\begin{aligned} P_I(F, m, \tau, \mathcal{K}, \alpha) &:= \text{Prob}[T_m(z) \geq \tau | H_0(m)], \\ P_{II}(F, m, \tau, \mathcal{K}, \alpha) &:= \text{Prob}[T_m(z) < \tau | H_1(m)]. \end{aligned}$$

We will characterize the *worst-case* error probabilities over all possible coalitions and users.

We first define the probability of a “false alarm”:

$$P_{fa}(F, \tau, \mathcal{K}, \alpha) := \max_{m \notin \mathcal{K}} P_I(F, m, \tau, \mathcal{K}, \alpha). \quad (7)$$

This is the probability of wrongly accusing the innocent user who looks most guilty. Equivalently, this is the probability of accusing at least one innocent user. The worst-case type I error probability is given by

$$P_I(F, \tau, \alpha) := \max_{\mathcal{K}} P_{fa}(F, \tau, \mathcal{K}, \alpha). \quad (8)$$

Next, consider the probability of a “miss”:

$$P_m(F, \tau, \mathcal{K}, \alpha) := \min_{m \in \mathcal{K}} P_{II}(F, m, \tau, \mathcal{K}, \alpha).$$

This is the probability of not accusing the most vulnerable guilty user. Equivalently, this is the probability of not detecting any colluders. Note that this event is the opposite of detecting at least one colluder:

$$P_d(F, \tau, \mathcal{K}, \alpha) := 1 - P_m(F, \tau, \mathcal{K}, \alpha). \quad (9)$$

The worst-case type II error probability is given by

$$P_{II}(F, \tau, \alpha) := \max_{\mathcal{K}} P_m(F, \tau, \mathcal{K}, \alpha). \quad (10)$$

The *worst-case* error probability is the maximum of the two error probabilities (8) and (10):

$$P_e(F, \tau, \alpha) := \max \{P_I(F, \tau, \alpha), P_{II}(F, \tau, \alpha)\}.$$

The threshold parameter τ can be varied to minimize this quantity, yielding the minmax error probability:

$$P_{\min\max}(F, \alpha) := \min_{\tau} P_e(F, \tau, \alpha). \quad (11)$$

In Section IV, we will analyze these error probabilities. We will also investigate average-case performance using simulations in Section V.

D. Geometric figure of merit

Related to the error probabilities is an important geometric figure of merit. For each user m , consider the distance between two types of potential collusions: those of which m is a member, and those of which m is not. Intuitively, if every noiseless fingerprint combination involving m is distant from every fingerprint combination not involving m , then even with moderate noise, there should be little ambiguity as to whether the m th user was involved or not.

To make this precise, for each user m , we define the “guilty” and “not guilty” sets of noiseless fingerprint combinations:

$$\begin{aligned} \mathcal{G}_m^{(K)} &:= \left\{ \frac{1}{|\mathcal{K}|} \sum_{k \in \mathcal{K}} f_k : m \in \mathcal{K} \subseteq \{1, \dots, M\}, |\mathcal{K}| \leq K \right\}, \\ -\mathcal{G}_m^{(K)} &:= \left\{ \frac{1}{|\mathcal{K}|} \sum_{k \in \mathcal{K}} f_k : m \notin \mathcal{K} \subseteq \{1, \dots, M\}, |\mathcal{K}| \leq K \right\}. \end{aligned}$$

In words, $\mathcal{G}_m^{(K)}$ is the set of size- K fingerprint combinations of equal weights ($\alpha_k = \frac{1}{K}$) which include m , while $-\mathcal{G}_m^{(K)}$ is the set of combinations which do not include m . Note that

in our setup (3), the α_k 's were arbitrary nonnegative values bounded by one. We will show in Section IV that the best attack from the collusion's perspective uses equal weights ($\alpha_k = \frac{1}{K}$) so that no single colluder is particularly vulnerable. For this reason, we use equal weights to obtain bounds on the distance between these two sets, which we define to be

$$\text{dist}(\mathcal{G}_m^{(K)}, -\mathcal{G}_m^{(K)}) := \min\{\|x - y\|_2 : x \in \mathcal{G}_m^{(K)}, y \in -\mathcal{G}_m^{(K)}\}. \quad (12)$$

In Section III, we find a lower bound on this distance.

Another related parameter is the *worst-case coherence*, which is the largest inner product between any two distinct fingerprints:

$$\mu := \max_{i \neq j} |\langle f_i, f_j \rangle|. \quad (13)$$

Intuitively, we want this value to be small, as this would correspond to having the fingerprints spaced apart. In the following section, we discuss a way of designing fingerprints which we will later evaluate using the above criteria.

III. ETF FINGERPRINT DESIGN

In this section, we introduce a fingerprint design based on equiangular tight frames (ETFs). We will discuss some important properties of ETFs and use them to determine a lower bound on the distance (12) from our geometric figure of merit. But first, we consider the fingerprinting problem from a compressed sensing viewpoint.

A. Compressed sensing viewpoint

We wish to design fingerprints in a way that will enable us to identify the small group of users who take part in a collusion. To do this, we consider a matrix-vector formulation of the attack (3) without noise, i.e., with $\epsilon = 0$. Specifically, let F denote the $N \times M$ matrix whose columns are the fingerprints $\{f_m\}_{m=1}^M$, and let the $M \times 1$ vector α denote the weights used in the collusion's linear average. Note that α_m is zero if user m is innocent, otherwise it's given by the corresponding coefficient in (3). This gives

$$z = F\alpha.$$

Using this representation, the detection problem can be interpreted from a compressed sensing perspective. Namely, α is a K -sparse vector that we wish to recover. Under certain conditions on the matrix F , we may "sense" this vector with $N < M$ measurements in such a way that the K -sparse vector is recoverable:

Theorem 1 ([36]). *Suppose an $N \times M$ matrix F satisfies the restricted isometry property (RIP):*

$$(1 - \delta_{2K})\|\alpha\|_2^2 \leq \|F\alpha\|_2^2 \leq (1 + \delta_{2K})\|\alpha\|_2^2 \quad (14)$$

for every $2K$ -sparse vector $\alpha \in \mathbb{R}^M$, where $\delta_{2K} \leq \sqrt{2} - 1$. Then for every K -sparse vector $\alpha \in \mathbb{R}^M$,

$$\alpha = \arg \min \|\hat{\alpha}\|_1 \text{ subject to } F\hat{\alpha} = F\alpha.$$

Thus, if F satisfies RIP (14), we can recover the K -sparse vector $\alpha \in \mathbb{R}^M$ by linear programming. For the attack model

with adversarial noise, $z = F\alpha + \epsilon$, if F satisfies RIP (14), linear programming will still produce an estimate $\hat{\alpha}$ of the sparse vector [37]. However, the distance between $\hat{\alpha}$ and α will be on the order of 10 times the size of the error ϵ . Due to potential legal ramifications of false accusations, this order of error is not tolerable. Note that these methods (both for the noiseless and noisy cases) recover the entire vector α ; equivalently, they identify the entire collusion. That said, we will investigate RIP matrices for fingerprint design, but to minimize false accusations, we will use focused detection (6) to identify colluders.

B. Geometric figure of merit for RIP matrices

We now investigate how well RIP matrices perform with respect to our geometric figure of merit. Without loss of generality, we assume the fingerprints are unit norm; since they have equal energy γ^2 , the fingerprint combination z can be normalized by γ before the detection phase. With this in mind, we have the following a lower bound on the distance (12) between the "guilty" and "not guilty" sets corresponding to any user m :

Theorem 2. *Suppose fingerprints $F = [f_1, \dots, f_M]$ satisfy the restricted isometry property (14). Then*

$$\text{dist}(\mathcal{G}_m^{(K)}, -\mathcal{G}_m^{(K)}) \geq \sqrt{\frac{1 - \delta_{2K}}{K(K-1)}}. \quad (15)$$

Proof: Take $\mathcal{K}, \mathcal{K}' \subseteq \{1, \dots, M\}$ such that $|\mathcal{K}|, |\mathcal{K}'| \leq K$ and $m \in \mathcal{K} \setminus \mathcal{K}'$. Then the left-hand inequality of the restricted isometry property (14) gives

$$\begin{aligned} & \left\| \frac{1}{|\mathcal{K}|} \sum_{m \in \mathcal{K}} f_m - \frac{1}{|\mathcal{K}'|} \sum_{m \in \mathcal{K}'} f_m \right\|^2 \\ &= \left\| \left(\frac{1}{|\mathcal{K}|} - \frac{1}{|\mathcal{K}'|} \right) \sum_{m \in \mathcal{K} \cap \mathcal{K}'} f_m + \frac{1}{|\mathcal{K}|} \sum_{m \in \mathcal{K} \setminus \mathcal{K}'} f_m - \frac{1}{|\mathcal{K}'|} \sum_{m \in \mathcal{K}' \setminus \mathcal{K}} f_m \right\|^2 \\ &\geq (1 - \delta_{|\mathcal{K} \cup \mathcal{K}'|}) \left(|\mathcal{K} \cap \mathcal{K}'| \left(\frac{1}{|\mathcal{K}|} - \frac{1}{|\mathcal{K}'|} \right)^2 + \frac{|\mathcal{K} \setminus \mathcal{K}'|}{|\mathcal{K}|^2} + \frac{|\mathcal{K}' \setminus \mathcal{K}|}{|\mathcal{K}'|^2} \right) \\ &= \frac{1 - \delta_{|\mathcal{K} \cup \mathcal{K}'|}}{|\mathcal{K}| |\mathcal{K}'|} \left(|\mathcal{K}| + |\mathcal{K}'| - 2|\mathcal{K} \cap \mathcal{K}'| \right). \end{aligned} \quad (16)$$

For a fixed $|\mathcal{K}|$, we will find a lower bound for

$$\frac{1}{|\mathcal{K}|} \left(|\mathcal{K}| + |\mathcal{K}'| - 2|\mathcal{K} \cap \mathcal{K}'| \right) = 1 + \frac{|\mathcal{K}| - 2|\mathcal{K} \cap \mathcal{K}'|}{|\mathcal{K}'|}. \quad (17)$$

Since we can have $|\mathcal{K} \cap \mathcal{K}'| > \frac{|\mathcal{K}|}{2}$, we know $\frac{|\mathcal{K}| - 2|\mathcal{K} \cap \mathcal{K}'|}{|\mathcal{K}'|} < 0$ when (17) is minimized. That said, $|\mathcal{K}'|$ must be as small as possible, that is, $|\mathcal{K}'| = |\mathcal{K} \cap \mathcal{K}'|$. Thus, when (17) is minimized, we must have

$$\frac{1}{|\mathcal{K}|} \left(|\mathcal{K}| + |\mathcal{K}'| - 2|\mathcal{K} \cap \mathcal{K}'| \right) = \frac{|\mathcal{K}|}{|\mathcal{K} \cap \mathcal{K}'|} - 1,$$

i.e., $|\mathcal{K} \cap \mathcal{K}'|$ must be as large as possible. Since $m \in \mathcal{K} \setminus \mathcal{K}'$, we have $|\mathcal{K} \cap \mathcal{K}'| \leq |\mathcal{K}| - 1$. Therefore,

$$\frac{1}{|\mathcal{K}|} \left(|\mathcal{K}| + |\mathcal{K}'| - 2|\mathcal{K} \cap \mathcal{K}'| \right) \geq \frac{1}{|\mathcal{K}| - 1}. \quad (18)$$

Substituting (18) into (16) gives

$$\left\| \frac{1}{|\mathcal{K}|} \sum_{m \in \mathcal{K}} f_m - \frac{1}{|\mathcal{K}'|} \sum_{m \in \mathcal{K}'} f_m \right\|^2 \geq \frac{1 - \delta_{|\mathcal{K} \cup \mathcal{K}'|}}{|\mathcal{K}|(|\mathcal{K}| - 1)} \geq \frac{1 - \delta_{2K}}{K(K - 1)}.$$

Since this bound holds for every m , \mathcal{K} and \mathcal{K}' with $m \in \mathcal{K} \setminus \mathcal{K}'$, we have (15). ■

Note that (15) depends on δ_{2K} , and so it's natural to ask how small δ_{2K} can be for a given fingerprint design. Also, which matrices even satisfy RIP (14)? In general, we say F is (K, δ) -RIP if for every K -sparse vector x ,

$$(1 - \delta)\|x\|_2^2 \leq \|Fx\|_2^2 \leq (1 + \delta)\|x\|_2^2.$$

Proving that a given matrix F is (K, δ) -RIP involves calculating eigenvalues of all size- K submatrices, which is computationally difficult for large matrices. The following lemma makes this explicit:

Lemma 3. *The smallest δ for which F is (K, δ) -RIP is*

$$\delta_{\min} := \max_{\substack{\mathcal{K} \subseteq \{1, \dots, M\} \\ |\mathcal{K}|=K}} \|F_{\mathcal{K}}^* F_{\mathcal{K}} - \mathbf{I}_K\|_2. \quad (19)$$

Proof: We first note that F being (K, δ) -RIP trivially implies that F is $(K, \delta + \epsilon)$ -RIP for every $\epsilon > 0$. It therefore suffices to show that (i) F is (K, δ_{\min}) -RIP, and (ii) F is not (K, δ) -RIP for any $\delta < \delta_{\min}$. To this end, pick some K -sparse vector x . To prove (i), we need to show that

$$(1 - \delta_{\min})\|x\|^2 \leq \|Fx\|^2 \leq (1 + \delta_{\min})\|x\|^2. \quad (20)$$

Let $\mathcal{K} \subseteq \{1, \dots, M\}$ be the size- K support of x , and let $x_{\mathcal{K}}$ be the corresponding subvector. Then rearranging (20) gives

$$\delta_{\min} \geq \left| \frac{\|Fx\|^2}{\|x\|^2} - 1 \right| = \left| \frac{\langle F_{\mathcal{K}} x_{\mathcal{K}}, F_{\mathcal{K}} x_{\mathcal{K}} \rangle - \langle x_{\mathcal{K}}, x_{\mathcal{K}} \rangle}{\|x_{\mathcal{K}}\|^2} \right| = \left| \left\langle \frac{x_{\mathcal{K}}}{\|x_{\mathcal{K}}\|}, (F_{\mathcal{K}}^* F_{\mathcal{K}} - \mathbf{I}_K) \frac{x_{\mathcal{K}}}{\|x_{\mathcal{K}}\|} \right\rangle \right|. \quad (21)$$

Since, by definition, δ_{\min} maximizes (21) over all supports \mathcal{K} and entry values $x_{\mathcal{K}}$, the inequality necessarily holds; that is, F is necessarily (K, δ_{\min}) -RIP. Furthermore, equality is achieved by the support \mathcal{K} which maximizes (19) and the eigenvector $x_{\mathcal{K}}$ corresponding to the largest eigenvalue of $F_{\mathcal{K}}^* F_{\mathcal{K}} - \mathbf{I}_K$; this proves (ii). ■

Since calculating eigenvalues for all size- K submatrices is computationally difficult, the Gershgorin circle theorem is often used to obtain a coarse bound in demonstrating RIP:

Theorem 4 (Gershgorin circle theorem [38]). *Take a $K \times K$ matrix A . Then for each eigenvalue λ of A , there exists an $i \in \{1, \dots, K\}$ such that λ lies in the complex disk centered at A_{ii} of radius $\sum_{j \neq i} |A_{ij}|$.*

This leads to the following well-known bound on δ_{2K} in terms of worst-case coherence:

Lemma 5. *Given a matrix F with unit-norm columns, then*

$$\delta_{2K} \leq (2K - 1)\mu, \quad (22)$$

where μ is the worst-case coherence (13).

Proof: From Lemma 3, the optimal δ_{2K} is

$$\delta_{2K} := \max_{\substack{\mathcal{K} \subseteq \{1, \dots, M\} \\ |\mathcal{K}|=2K}} \|F_{\mathcal{K}}^* F_{\mathcal{K}} - \mathbf{I}_K\|_2.$$

In words, given an ensemble of fingerprints, we consider each subcollection of size $2K$, subtract the identity from their $2K \times 2K$ Gram matrix and calculate the largest eigenvalue; the largest eigenvalue we find over all subcollections is what we call δ_{2K} . Since the fingerprints have unit norm, every Gram matrix of $2K$ fingerprints will have ones on the diagonal, and the absolute values of the off-diagonal entries will be no more than μ . Therefore, by applying the Gershgorin circle theorem (Theorem 4) to δ_{2K} , we obtain (22). ■

Combining Theorem 2 and Lemma 5 yields a coherence-based lower bound on the distance between the “guilty” and “not guilty” sets corresponding to any user m :

Theorem 6. *Suppose fingerprints $F = [f_1, \dots, f_M]$ are unit-norm with worst-case coherence μ . Then*

$$\text{dist}(\mathcal{G}_m^{(K)}, -\mathcal{G}_m^{(K)}) \geq \sqrt{\frac{1 - (2K - 1)\mu}{K(K - 1)}}. \quad (23)$$

We would like μ to be small, so that the lower bound (23) is as large as possible. But for a fixed N and M , the worst-case coherence of unit-norm fingerprints cannot be arbitrarily small; it necessarily satisfies the Welch bound [39]:

$$\mu \geq \sqrt{\frac{M - N}{N(M - 1)}}.$$

Equality in the Welch bound occurs precisely in the well-studied case where the fingerprints form an equiangular tight frame (ETF). An *equiangular tight frame* is a $N \times M$ matrix which has orthogonal rows of equal norm and unit-norm columns whose inner products have equal magnitude [40].

One type of ETF has already been proposed for fingerprint design: the simplex [20]. The simplex is an ETF with $M = N + 1$ and $\mu = \frac{1}{N}$. In fact, [20] gives a derivation for the value of the distance (12) in this case:

$$\text{dist}(\mathcal{G}_m^{(K)}, -\mathcal{G}_m^{(K)}) = \sqrt{\frac{1}{K(K - 1)} \frac{M}{M - 1}}. \quad (24)$$

The bound (23) is lower than (24) by a factor of $\sqrt{1 - \frac{2K}{N+1}}$, and for practical cases in which $K \ll N$, they are particularly close. Overall, ETF fingerprint design is a natural generalization of the provably optimal simplex design of [20].

C. Existence of RIP matrices

We now discuss the existence of RIP matrices as well as the construction of a specific class of ETFs. Given the RIP definition (14), it might not be clear how many matrices satisfy this condition. Surprisingly, RIP matrices are abundant; for any δ_{2K} , there exists a constant C such that if $N \geq CK \log(M/K)$, then matrices of Gaussian or Bernoulli (± 1) entries satisfy RIP with high probability. However, as mentioned in the previous section, checking if a given matrix satisfies RIP is computationally difficult.

Fortunately, there are some deterministic constructions of RIP matrices, such as [41], [42]. However the performance, measured by how large K can be, is not as good as the random constructions; the random constructions only require

$$F = \frac{1}{\sqrt{3}} \begin{bmatrix} + & - & + & - & + & - & + & - & & & & & & & & & \\ + & + & - & - & & & & & + & - & + & - & & & & & \\ + & - & - & + & & & & & & & & & + & - & + & - & \\ & & & & + & + & - & - & + & + & - & - & & & & & \\ & & & & + & - & - & + & & & & & + & + & - & - & \\ & & & & & & & + & - & - & + & + & - & - & + & \end{bmatrix}.$$

Fig. 2. The equiangular tight frame constructed in Example 1.

$N = \Omega(K \log^a M)$, while the deterministic constructions require $N = \Omega(K^2)$. The specific class of ETFs additionally requires $M \leq N^2$ [40].

Whether N scales as K versus K^2 is an important distinction between random and deterministic RIP matrices in the compressed sensing community. However, this difference offers no advantage for fingerprinting. Ergun et al. showed that for any fingerprinting system, a collusion of $K = O(\sqrt{N/\ln N})$ is sufficient to overcome the fingerprints [18]. This means with such a K , the detector cannot, with high probability, identify any attacker without incurring a significant false-alarm probability P_{fa} (7). This constraint is more restrictive than deterministic ETF constructions, as $\sqrt{N/\ln N} < \sqrt{N} \ll N/\log^a M$. Consequently, random RIP constructions are no better for fingerprint design than deterministic constructions.

Ergun's bound indicates that with a large enough K , colluders can overcome any fingerprinting system and render our detector unreliable. We now show that if K is sufficiently large, the colluders can *exactly* recover the original signal s :

Lemma 7. *Suppose the real equal-norm fingerprints $\{f_k\}_{k \in \mathcal{K}}$ do not lie in a common hyperplane. Then s is the unique minimizer of*

$$g(x) := \sum_{k \in \mathcal{K}} \left(\|x - (s + f_k)\|^2 - \frac{1}{|\mathcal{K}|} \sum_{k' \in \mathcal{K}} \|x - (s + f_{k'})\|^2 \right)^2.$$

Proof: Note that $g(x) \geq 0$, with equality precisely when $\|x - (s + f_k)\|^2$ is constant over $k \in \mathcal{K}$. Since the f_k 's have equal norm, $g(s) = 0$. To show that this minimum is unique, suppose $g(x) = 0$ for some $x \neq s$. This implies that the values $\|x - (s + f_k)\|^2$ are constant, with each being equal to their average. Moreover, since

$$\|x - (s + f_k)\|^2 = \|x - s\|^2 - 2\langle x - s, f_k \rangle + \|f_k\|^2,$$

we have that $\langle x - s, f_k \rangle$ is constant over $k \in \mathcal{K}$, contradicting the assumption that the fingerprints $\{f_k\}_{k \in \mathcal{K}}$ do not lie in a common hyperplane. ■

D. Construction of ETFs

Having established that deterministic RIP constructions are just as good as random constructions for fingerprint design, we now consider a particular method for constructing ETFs. Note that ETFs are notoriously difficult to construct in general, but a relatively simple approach was recently introduced in [42]. The approach uses a tensor-like combination of a Steiner system's adjacency matrix and a regular simplex, and it is

general enough to construct infinite families of ETFs. We illustrate the construction with an example:

Example 1 (see [42]). *To construct an ETF, we will use a simple class of Steiner systems, namely $(2, 2, v)$ -Steiner systems, and Hadamard matrices with v rows. In particular, $(2, 2, v)$ -Steiner systems can be thought of as all possible pairs of a v -element set [43], while a Hadamard matrix is a square matrix of ± 1 's with orthogonal rows [43].*

In this example, we take $v = 4$. The adjacency matrix of the $(2, 2, 4)$ -Steiner system has $\binom{4}{2} = 6$ rows, each indicating a distinct pair from a size-4 set:

$$A = \begin{bmatrix} + & + & & & & \\ + & & + & & & \\ + & & & + & & \\ & + & + & & & \\ & + & & + & & \\ & & + & + & & \end{bmatrix}. \quad (25)$$

To be clear, we use “+/-” to represent ± 1 and an empty space to represent a 0 value. Also, one example of a 4×4 Hadamard matrix is

$$H = \begin{bmatrix} + & + & + & + \\ + & - & + & - \\ + & + & - & - \\ + & - & - & + \end{bmatrix}. \quad (26)$$

We now build an ETF by replacing each of the “+” entries in the adjacency matrix (25) with a row from the Hadamard matrix (26). This is done in such a way that for each column of A , distinct rows of H are used; in this example, we use the second, third, and fourth rows of H . After performing this procedure, we normalize the columns, and the result is a real ETF F of dimension $N = 6$ with $M = 16$ fingerprints (see Figure 2).

We will use this method of ETF construction for simulations in Section V. Specifically, our simulations will let the total number of fingerprints M range from $2N$ to $7N$, a significant increase from the conventional $N + 1$ simplex and N orthogonal fingerprints. Note, however, that unlike simplex and orthogonal fingerprints, there are limitations to the dimensions that admit real ETFs. For instance, the size of a Hadamard matrix is necessarily a multiple of 4, and the existence of Steiner systems is rather sporadic. Fortunately, identifying Steiner systems is an active area of research, with multiple infinite families already characterized and alternative construction methods developed [42], [44].

IV. ERROR ANALYSIS

A. Analysis of type I and type II errors

We now investigate the worst case errors involved with using ETF fingerprint design and focused correlation detection under linear averaging attacks. Recall that the worst-case type I error probability (8) is the probability of falsely accusing the most guilty-looking innocent user. Also recall that the worst-case type II error probability (10) is the probability of not accusing the most vulnerable guilty user.

Theorem 8. Suppose the fingerprints $F = \{f_m\}_{m=1}^M$ form an equiangular tight frame. Then the worst-case type I and type II error probabilities, (8) and (10), satisfy

$$\begin{aligned} P_I(F, \tau, \alpha) &\leq Q\left[\frac{\gamma}{\sigma}(\tau - \mu)\right], \\ P_{II}(F, \tau, \alpha) &\leq Q\left[\frac{\gamma}{\sigma}\left((1 + \mu)\max_{k \in \mathcal{K}} \alpha_k - \mu\right) - \tau\right], \end{aligned}$$

where $Q(x) := \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-u^2/2} du$ and $\mu = \sqrt{\frac{M-N}{N(M-1)}}$.

Proof: Under hypothesis $H_0(m)$, the test statistic for the detector (6) is

$$\begin{aligned} T_z(m) &= \frac{1}{\gamma^2} \left\langle \sum_{n \in \mathcal{K}} \alpha_n f_n + \epsilon, f_m \right\rangle \\ &= \sum_{n \in \mathcal{K}} \alpha_n (\pm \mu) + \epsilon', \end{aligned}$$

where ϵ' is the projection of noise ϵ/γ onto the normed vector f_m/γ , and, due to symmetry of the variance of ϵ in all dimensions, $\epsilon' \sim \mathcal{N}(0, \sigma^2/\gamma^2)$. Thus, under hypothesis $H_0(m)$, $T_z(m) \sim \mathcal{N}(\sum_{n \in \mathcal{K}} \alpha_n (\pm \mu), \sigma^2/\gamma^2)$. We can subtract the mean and divide by the standard deviation to obtain:

$$\begin{aligned} \text{Prob}[T_m(z) \geq \tau | H_0(m)] &= Q\left(\frac{\gamma}{\sigma} \left[\tau - \left(\sum_{n \in \mathcal{K}} \alpha_n (\pm \mu) \right) \right]\right) \\ &\leq Q\left(\frac{\gamma}{\sigma}(\tau - \mu)\right). \end{aligned} \quad (27)$$

Note that $Q(x)$ is a decreasing function. The bound (27) is obtained by setting all of the coefficients of the coherence to be positive.

Likewise, under hypothesis $H_1(m)$, the test statistic is

$$\begin{aligned} T_z(m) &= \frac{1}{\gamma^2} \left\langle \sum_{n \in \mathcal{K}} \alpha_n f_n + \epsilon, f_m \right\rangle \\ &= \alpha_m + \sum_{n \in \mathcal{K} \setminus \{m\}} \alpha_n (\pm \mu) + \epsilon'. \end{aligned}$$

Thus, under hypothesis $H_1(m)$,

$$T_z(m) \sim \mathcal{N}\left(\alpha_m + \sum_{n \in \mathcal{K} \setminus \{m\}} \alpha_n (\pm \mu), \frac{\sigma^2}{\gamma^2}\right).$$

Since $1 - Q(x) = Q(-x)$, the type II error probability can be bounded as

$$\begin{aligned} \text{Prob}[T_m(z) < \tau] &= Q\left(-\frac{\gamma}{\sigma} \left[\tau - \left(\alpha_m + \sum_{n \in \mathcal{K} \setminus \{m\}} \alpha_n (\pm \mu) \right) \right]\right) \\ &\leq Q\left(\frac{\gamma}{\sigma}([\alpha_m(1 + \mu) - \mu] - \tau)\right). \end{aligned}$$

We can now evaluate the worst-case errors (8) and (10). For type I errors, with $m \notin \mathcal{K}$:

$$\begin{aligned} P_I(F, \tau, \alpha) &= \max_{\mathcal{K}} \max_{m \notin \mathcal{K}} P_I(F, m, \tau, \mathcal{K}, \alpha) \\ &= \max_{\mathcal{K}} \max_{m \notin \mathcal{K}} \text{Prob}[T_m(z) \geq \tau | H_0(m)] \\ &\leq \max_{\mathcal{K}} \max_{m \notin \mathcal{K}} Q\left(\frac{\gamma}{\sigma}(\tau - \mu)\right) \\ &= Q\left(\frac{\gamma}{\sigma}(\tau - \mu)\right). \end{aligned}$$

For type II errors,

$$\begin{aligned} P_{II}(F, \tau, \alpha) &= \max_{\mathcal{K}} \min_{m \in \mathcal{K}} P_{II}(F, m, \tau, \mathcal{K}, \alpha) \\ &= \max_{\mathcal{K}} \min_{m \in \mathcal{K}} \text{Prob}[T_m(z) < \tau] \\ &\leq \max_{\mathcal{K}} \min_{m \in \mathcal{K}} Q\left(\frac{\gamma}{\sigma}([\alpha_m(1 + \mu) - \mu] - \tau)\right) \\ &= Q\left(\frac{\gamma}{\sigma}([\max_{m \in \mathcal{K}} \alpha_m(1 + \mu) - \mu] - \tau)\right). \end{aligned}$$

The last equation follows since once the α 's are fixed, the actual coalition \mathcal{K} does not matter. ■

We can further maximize over all possible weightings α :

$$\begin{aligned} P_I(F, \tau) &= \max_{\alpha} P_I(F, \tau, \alpha) \\ &\leq Q\left(\frac{\gamma}{\sigma}(\tau - \mu)\right) \\ P_{II}(F, \tau) &= \max_{\alpha} P_{II}(F, \tau, \alpha) \\ &\leq \max_{\alpha} Q\left(\frac{\gamma}{\sigma}([\max_{m \in \mathcal{K}} \alpha_m(1 + \mu) - \mu] - \tau)\right) \\ &= Q\left(\frac{\gamma}{\sigma}([\min_{\alpha} \max_{m \in \mathcal{K}} \alpha_m(1 + \mu) - \mu] - \tau)\right) \\ &= Q\left(\frac{\gamma}{\sigma}([\frac{1}{K}(1 + \mu) - \mu] - \tau)\right). \end{aligned}$$

Thus, the vector α which minimizes the value of its maximum element is the uniform weight vector. This motivates the attackers' use of a uniformly weighted coefficient vector α to maximize the probability that none of the members will be caught.

B. Minimax error analysis

From a detection standpoint, an important criterion is the minimax error probability (11). The threshold τ trades off type I and type II errors. Thus, there is a value of τ , denoted by τ^* , which minimizes the maximum of the probabilities of the two error types. Since the bound for the type I error probability is independent of α , and the bound for the type II error probability is maximized with a uniform weighting, assume this to be the case.

Theorem 9. *The minmax probability of error (11) can be bounded as:*

$$Q\left(\frac{d_{\text{low}}^*}{2}\right) \leq P_{\text{minmax}}(F, \alpha) \leq Q\left(\frac{d_{\text{up}}^*}{2}\right), \quad (28)$$

where

$$d_{\text{low}}^* := \frac{\sqrt{\frac{M}{M-1}} \sqrt{ND_f}}{\sigma \sqrt{K(K-1)}}$$

$$d_{\text{up}}^* := \frac{\sqrt{ND_f}}{\sigma K} \left(1 - (2K-1)\mu\right).$$

Note that for orthogonal and simplex fingerprints, the minmax errors are both of the form

$$P_{\text{minmax}}(F, \alpha) = Q\left(\frac{d^*(K)}{2}\right).$$

For orthogonal fingerprints [20],

$$d^*(K) = \frac{\sqrt{ND_f}}{\sigma K},$$

which is better than $d_{\text{low}}^*(K)$ (the Q function is decreasing). For simplex fingerprints, $d^*(K)$ is slightly better than both [20]:

$$d^*(K) = \frac{\sqrt{ND_f}}{\sigma K} \frac{M}{M-1}.$$

Proof: The lower bound is the sphere packing lower bound [24]. For the upper bound,

$$P_e(F, \tau, \alpha) = \max\{P_I(F, \tau, \alpha), P_{II}(F, \tau, \alpha)\}$$

$$\leq \max\left\{Q\left(\frac{\gamma}{\sigma}(\tau - \mu)\right), Q\left(\frac{\gamma}{\sigma}\left(\left\lceil \frac{1}{K}(1+\mu) - \mu \right\rceil - \tau\right)\right)\right\}.$$

Since the test statistic $T_z(m)$ is normally distributed with the same variance under either of the hypotheses $H_0(m)$ and $H_1(m)$, the value of τ that minimizes this upper bound is the average of the means μ and $\frac{1+\mu}{K} - \mu$, namely $\tau^* := \frac{1+\mu}{2K}$. Using this τ^* and recalling (2), we have

$$P_{\text{minmax}}(F, \alpha) = \min_{\tau} P_e(F, \tau, \alpha) = P_e(F, \tau^*, \alpha)$$

$$\leq Q\left(\frac{\gamma}{\sigma}(\tau^* - \mu)\right)$$

$$= Q\left(\frac{\sqrt{ND_f}}{2\sigma K} \left(1 - (2K-1)\mu\right)\right)$$

$$= Q(d_{\text{up}}^*(K)/2).$$

Consider the regime where N is large, M grows linearly or faster than N , and WNR is constant (in particular 0, so $D_f = \sigma^2$). Then $\mu \approx 1/\sqrt{N}$,

$$d_{\text{low}}^* \approx \frac{\sqrt{N}}{K}, \quad \text{and} \quad d_{\text{up}}^* \approx \frac{\sqrt{N}}{K} \left(1 - \frac{2K}{\sqrt{N}}\right) = \frac{\sqrt{N}}{K} - 2.$$

If $K \ll \sqrt{N}$, then both bounds go to infinity so $P_{\text{minmax}}(F, \alpha) \rightarrow 0$. Also, the geometric figure of merit (12) then behaves as

$$\text{dist}(\mathcal{G}_m^{(K)}, -\mathcal{G}_m^{(K)}) \approx \frac{1}{K} \approx \sqrt{N} d_{\text{up}}^*.$$

If K is proportional to \sqrt{N} , then $P_{\text{minmax}}(F, \alpha)$ is bounded away from 0.

We can also compute the error exponent for this test:

$$e(F, \tau^*, \alpha) := -\lim_{N \rightarrow \infty} \frac{1}{N} \ln P_e(F, \tau^*, \alpha).$$

Corollary 10. *If $M \gg N \gg K^2$, then the error exponent is*

$$e(F, \tau^*, \alpha) = \frac{1}{8K^2}. \quad (29)$$

Proof: The proof follows by applying the asymptotic equality $\ln Q(t) \sim -\frac{t^2}{2}$ as $t \rightarrow \infty$ to the bounds in (28). The bounds are asymptotically equivalent. ■

Note that this error exponent is the same as in the simplex case [20]. As $K \rightarrow \infty$, the error exponent (29) goes to zero.

V. SIMULATIONS

In Section IV, the worst-case error probabilities were analyzed, where the worst case was over all collusions. Here we investigate average case behavior. We examine the probability of detecting at least one guilty user, P_d (9), as a function of K with the false-alarm P_{fa} (7) fixed below a threshold. The threshold can be interpreted as the (legally) allowable limit for probability of false accusation. We compare the average case performance of ETF fingerprints, simplex fingerprints [20], and orthogonal fingerprints [16] for four dimension sizes $N \in \{195, 651, 2667, 8128\}$. The ETF construction described in Example 1 was used. The results demonstrate that ETF fingerprints perform almost as well as both orthogonal and simplex fingerprints, while accommodating several times as many users. We now describe the design of the simulations.

A. Design

For a fixed signal dimension size N , ETF, orthogonal, and simplex fingerprints were created. The ETF fingerprint design was constructed using the method shown in Example 1. For $N = 195$, a $(2, 7, 91)$ -Steiner system was used [44], yielding $M = 1456$ fingerprints. For $N = 651$ and $N = 2667$, the Steiner systems were constructed using projective geometry, with $M = 2016$ and $M = 8128$ respectively [42]. For $N = 8128$, a $(2, 2, 2^7)$ -Steiner system was used, giving $M = 16,384$ fingerprints [42].

The orthogonal fingerprint design was constructed using an identity matrix. For actual embedding, the orthogonal fingerprints should be randomly rotated to ensure difficulty in removing them. One method to achieve this is to randomly generate i.i.d. Gaussian vectors in N dimensional space, orthogonalize them, then scale them to have the same energy [16]. For detection purposes, however, rotations of the basis vectors are inconsequential.

The simplex fingerprint design was constructed using the following method of sequentially fixing the values of the fingerprints in each dimension. The vectors of a regular simplex are equidistant from the center, having the same power, and have inner products equal to $-\frac{1}{N}$ [45]. Letting the first vector have a one in the first row and zeros in the other rows forces every other vector's first element to be $-\frac{1}{N}$ to satisfy the inner product constraint. For the second vector, choose its second

element to satisfy the power constraint and set the elements of the remaining rows to zero. Using the inner product constraint, we can find the value of the second element for all other vectors. Repeating these steps constructs a regular simplex.

Linear collusion attacks were simulated separately for the different designs and collusion sizes K . For each attack, K of the M fingerprints were randomly chosen and uniformly averaged. Next, an i.i.d. Gaussian noise vector was added with per-sample noise power $\sigma^2 = D_f$, corresponding to a WNR (4) of 0 dB [19].

The test statistics $T_z(m)$ (5) were then computed for each user m . For each threshold τ , it was determined whether there was a detection event (at least one colluder with $T_z(m) > \tau$) and/or a false alarm (at least one innocent user with $T_z(m) > \tau$). In total, 50,000 attacks were simulated, and the detection and false alarm counts were averaged. Then the minimal τ value was selected for which $P_{fa} \leq 10^{-3}$. This induced the corresponding P_d .

B. Results

We ran experiments with four different dimension sizes $N \in \{195, 651, 2667, 8128\}$. The noise level was kept at WNR = 0 dB. The value of K varied between 1 and sufficiently large values so P_d approached zero. Plots for the probability of detection P_d as a function of the size of the coalition K are shown in Figures 3–6 for $N \in \{195, 651, 2667, 8128\}$ respectively. The largest values of K for which at least one attacker can be caught with probability (nearly) one under the P_{fa} constraint are about 2, 5, 11, and 21 respectively. Overall, ETF fingerprints perform comparably to orthogonal and simplex fingerprints while accommodating many more users.

REFERENCES

- [1] D. Boneh and J. Shaw, "Collusion-secure fingerprinting for digital data," *Information Theory, IEEE Transactions on*, vol. 44, no. 5, pp. 1897–1905, 1998.
- [2] B. Chor, A. Fiat, M. Naor, and B. Pinkas, "Tracing traitors," *Information Theory, IEEE Transactions on*, vol. 46, no. 3, pp. 893–910, 2000.
- [3] H. Schaathun, "The boneh-shaw fingerprinting scheme is better than we thought," *Information Forensics and Security, IEEE Transactions on*, vol. 1, no. 2, pp. 248–255, 2006.
- [4] G. Tardos, "Optimal probabilistic fingerprint codes," *Journal of the ACM (JACM)*, vol. 55, no. 2, p. 10, 2008.
- [5] A. Somekh-Baruch and N. Merhav, "On the capacity game of private fingerprinting systems under collusion attacks," *Information Theory, IEEE Transactions on*, vol. 51, no. 3, pp. 884–899, 2005.
- [6] —, "Achievable error exponents for the private fingerprinting game," *Information Theory, IEEE Transactions on*, vol. 53, no. 5, pp. 1827–1838, 2007.
- [7] N. Ananthapadmanabhan, A. Barg, and I. Dumer, "On the fingerprinting capacity under the marking assumption," *Information Theory, IEEE Transactions on*, vol. 54, no. 6, pp. 2678–2689, 2008.
- [8] P. Moulin, "Universal fingerprinting: Capacity and random-coding exponents," *Arxiv preprint arXiv:0801.3837*, 2008.
- [9] A. Barg, G. Blakley, and G. Kabatiansky, "Digital fingerprinting codes: Problem statements, constructions, identification of traitors," *Information Theory, IEEE Transactions on*, vol. 49, no. 4, pp. 852–865, 2003.
- [10] S. Lin, M. Shahmohammadi, and H. El Gamal, "Fingerprinting with minimum distance decoding," *Information Forensics and Security, IEEE Transactions on*, vol. 4, no. 1, pp. 59–69, 2009.
- [11] M. Cheng and Y. Miao, "On anti-collusion codes and detection algorithms for multimedia fingerprinting," *Information Theory, IEEE Transactions on*, vol. 57, no. 7, pp. 4843–4851, 2011.
- [12] J. Cotrina-Navau and M. Fernández, "A family of asymptotically good binary fingerprinting codes," *Information Theory, IEEE Transactions on*, vol. 56, no. 10, pp. 5335–5343, 2010.
- [13] D. Boneh, A. Kiayias, and H. Montgomery, "Robust fingerprinting codes: a near optimal construction," in *Proceedings of the tenth annual ACM workshop on Digital rights management*. ACM, 2010, pp. 3–12.
- [14] H. Koga and Y. Minami, "A digital fingerprinting code based on a projective plane and its identifiability of all malicious users," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 94, no. 1, pp. 223–232, 2011.
- [15] W. Trappe, M. Wu, Z. Wang, and K. Liu, "Anti-collusion fingerprinting for multimedia," *Signal Processing, IEEE Transactions on*, vol. 51, no. 4, pp. 1069–1087, 2003.
- [16] I. Cox, J. Kilian, F. Leighton, and T. Shamoan, "Secure spread spectrum watermarking for multimedia," *Image Processing, IEEE Transactions on*, vol. 6, no. 12, pp. 1673–1687, 1997.
- [17] J. Kilian, F. Leighton, L. Matheson, T. Shamoan, R. Tarjan, and F. Zane, "Resistance of digital watermarks to collusive attacks," in *IEEE International Symposium on Information Theory*, 1998, pp. 271–271.
- [18] F. Ergun, J. Kilian, and R. Kumar, "A note on the limits of collusion-resistant watermarks," in *Advances in Cryptology, EUROCRYPT 99*. Springer, 1999, pp. 140–149.
- [19] Z. Wang, M. Wu, H. Zhao, W. Trappe, and K. Liu, "Anti-collusion forensics of multimedia fingerprinting using orthogonal modulation," *Image Processing, IEEE Transactions on*, vol. 14, no. 6, pp. 804–821, 2005.
- [20] N. Kiyavash, P. Moulin, and T. Kalker, "Regular simplex fingerprints and their optimality properties," *Information Forensics and Security, IEEE Transactions on*, vol. 4, no. 3, pp. 318–329, Sept. 2009.
- [21] N. Hayashi, M. Kuribayashi, and M. Morii, "Collusion-resistant fingerprinting scheme based on the cdma-technique," in *Proceedings of the Security 2nd international conference on Advances in information and computer security*. Springer-Verlag, 2007, pp. 28–43.
- [22] Z. Li and W. Trappe, "Collusion-resistant fingerprints from wbe sequence sets," in *Communications, 2005. ICC 2005. 2005 IEEE International Conference on*, vol. 2. IEEE, 2005, pp. 1336–1340.
- [23] J. Jourdas and P. Moulin, "High-rate random-like spherical fingerprinting codes with linear decoding complexity," *Information Forensics and Security, IEEE Transactions on*, vol. 4, no. 4, pp. 768–780, 2009.
- [24] N. Kiyavash and P. Moulin, "Sphere packing lower bound on fingerprinting error probability," in *Proceedings of SPIE*, vol. 6505, 2007.
- [25] O. Dalkilic, E. Ekrem, S. Varlik, and M. Mihcak, "A detection theoretic approach to digital fingerprinting with focused receivers under uniform linear averaging gaussian attacks," *Information Forensics and Security, IEEE Transactions on*, vol. 5, no. 4, pp. 658–669, 2010.
- [26] Y. Wang and P. Moulin, "Capacity and optimal collusion attack channels for gaussian fingerprinting games," vol. 6505, no. 1. SPIE, 2007, p. 65050J.
- [27] H. Ling, H. Feng, F. Zou, W. Yan, and Z. Lu, "A novel collusion attack strategy for digital fingerprinting," *Digital Watermarking*, pp. 224–238, 2011.
- [28] N. Kiyavash and P. Moulin, "Performance of orthogonal fingerprinting codes under worst-case noise," *Information Forensics and Security, IEEE Transactions on*, vol. 4, no. 3, pp. 293–301, 2009.
- [29] R. Pickholtz, D. Schilling, and L. Milstein, "Theory of spread-spectrum communications—a tutorial," *Communications, IEEE Transactions on*, vol. 30, no. 5, pp. 855–884, 1982.
- [30] I. Cox, M. Miller, and J. Bloom, *Digital watermarking*. Morgan Kaufmann Pub, 2002.
- [31] F. Hartung and M. Kutter, "Multimedia watermarking techniques," *Proceedings of the IEEE*, vol. 87, no. 7, pp. 1079–1107, 1999.
- [32] W. Dai, N. Kiyavash, and O. Milenkovic, "Spherical codes for sparse digital fingerprinting," in *Spring Central Meeting of the American Mathematical Society, AMS08, Special Session on Algebraic Aspects of Coding Theory*, 2008.
- [33] D. Varodayan and C. Pépin, "Collusion-aware traitor tracing in multimedia fingerprinting using sparse signal approximation," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*. IEEE, 2008, pp. 1645–1648.
- [34] H. Pham, W. Dai, and O. Milenkovic, "Compressive list-support recovery for colluder identification," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*. IEEE, pp. 4166–4169.
- [35] C. Colbourn, D. Horsley, and V. Syrotiuk, "Frameproof codes and compressive sensing," in *Communication, Control, and Computing (Allerton), 48th Allerton Conference on*. IEEE, 2010, pp. 985–990.

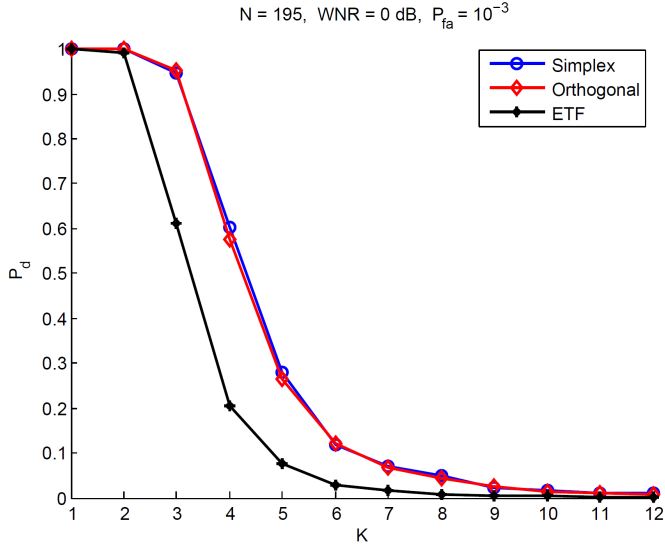


Fig. 3. A plot of the probability of detecting at least one colluder (P_d) as a function of the number of colluders (K). The threshold τ is picked to be the minimum threshold to fix $P_{fa} \leq 10^{-3}$. The WNR is 0 dB and $N = 195$. The (maximum) number of fingerprints were 195 for the orthogonal, 196 for the simplex, and 1456 for the ETF construction.

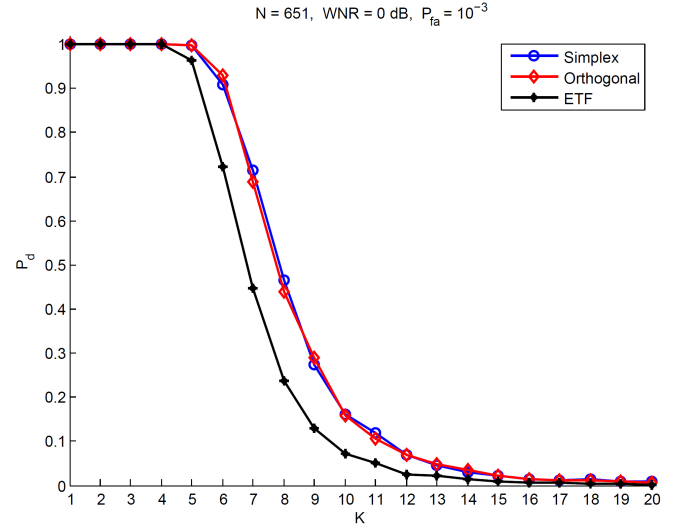


Fig. 4. A plot of the probability of detecting at least one colluder (P_d) as a function of the number of colluders (K). The threshold τ is picked to be the minimum threshold to fix $P_{fa} \leq 10^{-3}$. The WNR is 0 dB and $N = 651$. The (maximum) number of fingerprints were 651 for the orthogonal, 652 for the simplex, and 2016 for the ETF construction.

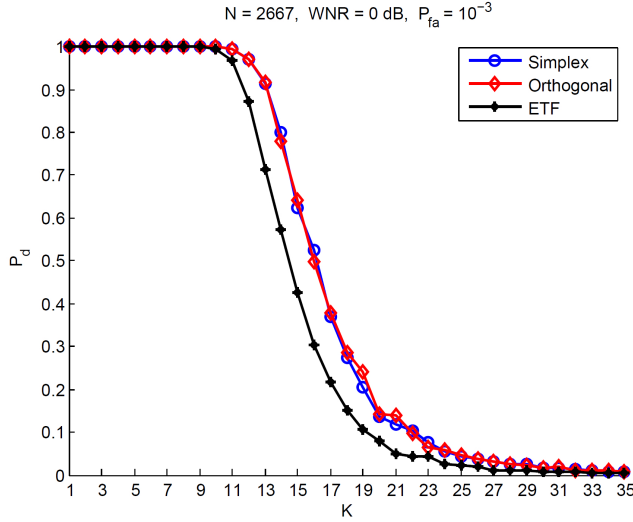


Fig. 5. A plot of the probability of detecting at least one colluder (P_d) as a function of the number of colluders (K). The threshold τ is picked to be the minimum threshold to fix $P_{fa} \leq 10^{-3}$. The WNR is 0 dB and $N = 2667$. The (maximum) number of fingerprints were 2667 for the orthogonal, 2668 for the simplex, and 8128 for the ETF construction.

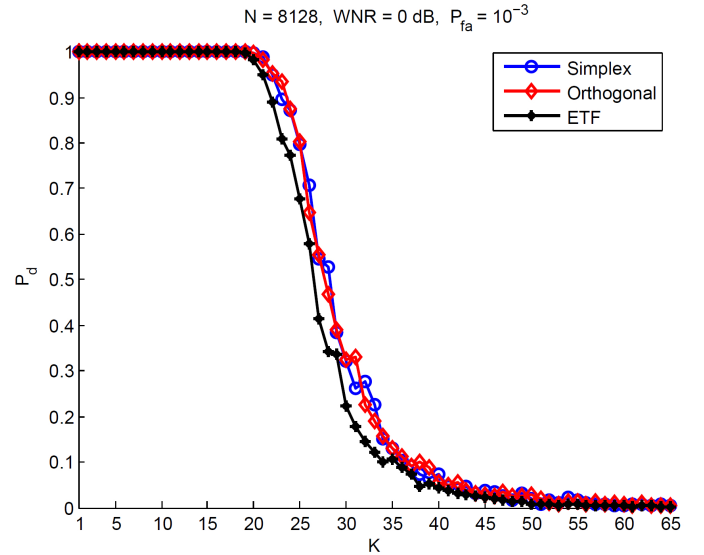


Fig. 6. A plot of the probability of detecting at least one colluder (P_d) as a function of the number of colluders (K). The threshold τ is picked to be the minimum threshold to fix $P_{fa} \leq 10^{-3}$. The WNR is 0 dB and $N = 8128$. The (maximum) number of fingerprints were 8128 for the orthogonal, 8129 for the simplex, and 16,384 for the ETF construction.

- [36] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.
- [37] E. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on pure and applied mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.
- [38] R. Varga, *Geršgorin and his circles*. Springer Verlag, 2004.
- [39] L. Welch, "Lower bounds on the maximum cross correlation of signals (corresp.)," *Information Theory, IEEE Transactions on*, vol. 20, no. 3, pp. 397–399, 1974.
- [40] T. Strohmer and R. W. Heath, "Grassmannian frames with applications to coding and communication," *Applied and Computational Harmonic Analysis*, vol. 14, no. 3, pp. 257 – 275, 2003.
- [41] R. DeVore, "Deterministic constructions of compressed sensing matrices," *Journal of Complexity*, vol. 23, no. 4-6, pp. 918–925, 2007.
- [42] M. Fickus, D. Mixon, and J. Tremain, "Steiner equiangular tight frames," *Linear Algebra Appl.*, 2011, to appear.
- [43] J. van Lint and R. Wilson, *A course in combinatorics*, 2nd ed. Cambridge University Press, 2001.
- [44] "Steiner systems," in *La Jolla Covering Repository*, 2011, <http://www.ccrwest.org/cover/steiner.html>.
- [45] J. Munkres, *Elements of algebraic topology*. Addison-Wesley Reading, MA, 1984, vol. 2.